DOCUMENT RESUME

ED 117 117 TM 003 922

AUTHOR Whitlock, Gerald H.
TITLE Establishing Passing Scores for Professional Licensing Examinations.
PUB DATE [Jul 74]
NOTE 6p.

EDRS PRICE MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS Certification; *Cutting Scores; Occupational Tests; *Predictive Validity; Professional Education; Scoring Formulas; *State Licensing Boards; *Test Validity
IDENTIFIERS Criterion Related Validity

ABSTRACT

Until data is obtained concerning the regression of job-performance on test-performance, the setting of passing scores on professional licensing and certification examinations will contain some degree of arbitrariness. Data from performance domains suggests that some tests have differential validity and an adverse impact on minority groups, disproportionately excluding them from professional practice on the absence of any data indicating criterion related validity of the licensing examinations. The measurement of content validity for these examinations demonstrates wide differences among professions and across states within the same professions. Some typical methods for setting passing scores are shown. The author introduces his own methodology for establishing passing scores and outlines its mechanics. Some of the unresolved questions of the method are discussed briefly. (SE)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED117117

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ESTABLISHING PASSING SCORES FOR PROFESSIONAL LICENSING EXAMINATIONS

Gerald H. Whitlock, Ph.D.
University of Tennessee

This paper is concerned with the problems associated with the setting of passing scores on professional licensing and certification examinations. Since empirical studies of the criterion related validity of these kinds of tests are practically non-existent, no useful data are available by which to relate successful and unsuccessful test performance to successful and unsuccessful professional practice. And until one obtains the data describing the regression of job-performance on test performance, the setting of cutting scores or passing levels will have an element of arbitrariness. However, data from other performance domains which suggest that some tests have differential validity for minority groups, along with the fact that in some states professional licensing examinations have been shown to have an adverse impact on the members of minority groups, together suggest that licensing examinations, even if they do have a degree of over-all validity, may in fact have differential validity for minority groups. This would then seem to cast the problem in terms of the EEOC guidelines for compliance with Title VII of the 1964 Civil Rights Act since minority group members are being excluded from the practice of their profession in disproportionate numbers in the absence of any data indicating criterion related validity of the licensing examinations for selection into the profession, and it is not known whether the scores assigned to minority group members have the same predictive properties, from the stand point of competence to practice, as those assigned to majority group members.

TM 003 922

Opinions differ as to whether it is feasible, or indeed possible, to conduct criterion related validity studies in any of the professions. The problem of acquiring meaningful measures of job performance is cited as the difficulty. Such a position is made more difficult to defend in view of the fact that performance evaluations are routinely conducted in the professions of engineering, accounting, teaching and law. Peterson has described a study in North Carolina where good reliability was obtained for ratings of performance of medical practitioners. Even so, it does not appear likely that any such studies are going to be conducted in the near future for the purpose of setting passing scores on licensing examinations, so consideration needs to be given to some alternatives which at least in some sense minimize arbitrariness.

Current practice is to rely on content validity of professional examinations. There are marked differences among the professions for measuring content validity and the quality of the procedures ranges from sloppy to excellent. Even in the same profession there may be wide differences in methods across states and with the same range of quality.

In any case the means are assumed to produce the desired ends, so the examinations are administered at prescribed times and candidates who do not reach pre-determined levels (a score of 70 on a 100 point scale seems to be favored) are not admitted to practice. In some cases, as in CPA exams, parts may be repeated and in others the entire exam must be repeated, e.g., law. The concern here is with determining the passing score. First are considered some typical methods for accomplishing this (remembering that the following methods are used in the absence of any criterion related validity data).

Use of a Norm Group

A. This method uses an appropriate norm group, and sets the passing score in terms of their performance. For example, one of the professions in Illinois admits candidates to the exam who are rather heterogeneous with respect to educational achievement. Thus they have found that passing everyone whose score is above one standard deviation below the mean passes about the "right" proportion of candidates (very few of the candidates with Ph.D's score that low).

B. Where only those candidates are admitted to the exams who have graduated from an accredited (by the profession) school, then one might go two standard deviations below the mean to set the passing score. This method assumes that a very small percentage of such graduates should fail.

Absolute Grading

In this method a predetermined passing score is set. Typically the score is 70 or 75, and anyone scoring below that is failed. This is the typical method when the exam is essay or problem solving, e.g., law, part of engineering, accounting, medicine. In some cases borderline cases are regraded, in others not. In some cases great care is taken to achieve reliability of grading, in others almost none. The writer knows of several states in which parts of the essay exams of a major profession are scored by only one individual. As might be expected, such a procedure results in considerable variations from year to year in the number of candidates admitted to the profession. There are some obvious ways to improve the reliability of such scores, e.g., multiple graders, use of model answers, and the like, and given the test reliability, it is then possible to be sure that with any desired probability the passing score actually used would not exclude someone whose true achievement was at the level of the

nominal passing score (e.g., if the nominal passing score was 70, then a passing score 1.645 standard errors of measurement below 70 would provide 95% confidence that failing candidates would not have been passed).

It can be seen that the major element of arbitrariness in all of the above methods is the setting of passing scores. Even the setting of particular confidence limits is arbitrary, although bolstered by tradition in some contexts. This arbitrariness will remain in the absence of empirical validation data. However, the writer believes that meanwhile for some of the professional exams and parts of others, the setting of passing scores can be put on firmer grounds, both logically and statistically. The basic assumption of the method proposed is that faculty of accredited professional schools in cooperation with the accreditors provide the proper professional instruction and maintain proper standards of accomplishment to insure that their graduates are competent to practice in the profession. The mechanisms of the method are as follows:

1. For all examinees for a particular examination administration who are recent graduates of an accredited school within the state, determine the regression equation for predicting examination scores from professional school Grade Point Average. ($Y = a + b(GPA)$)
2. Calculate the predicted examination score for the lowest GPA required to graduate. This is the nominal passing score.
3. Subtract 1.645 standard errors of estimate from the nominal passing score to arrive at the actual passing score. Thus one is 95% confident that an individual who failed should not have been passed.

This method applies where the examination is taken, or those parts of it, before or soon after graduation, e.g., bar exam. Other methods would need to be used for those parts of the exam which are given only after some number of years in practice, e.g., some certifying exams.

Using the above method insures (with high probability) that anyone who is admitted to practice has at least the competence of a graduate of an accredited school.

Also, in the event that the exam scores and GPA correlated differently for members of minority groups, the separate regression equations could be used.

Some Unresolved Questions

Many states have more than one accredited school in a particular profession so the question arises which one to use, since the equations will differ along with minimum GPA's for graduation. There would seem to be some merit in using the official state school to set the standard. However, other accredited schools may then suffer if they attract less able students. Or, it could be argued that the equation should be based on the school yielding the lowest passing score since this score does reflect a level of achievement that presumably meets accreditation standards. A compromise would be to use separate equations to determine the passing score for each school, and use their unweighted mean passing score as the official passing score.